

Collaborative Virtual Training with Physical and Communicative Autonomous Agents.

Thomas Lopez¹

Pierre Chevaillier²

Valrie Gouranton¹

Paul Evrard³

Florian Nouviale¹

Mukesh Barange²

Rozenn Bouville¹

Bruno Arnaldi¹

Abstract

Virtual agents are a real asset in Collaborative Virtual Environment for Training (CVET) as they can replace missing team members. Collaboration between such agents and users, however, is generally limited. We present here a whole integrated model of CVET focusing on the abstraction of the real or virtual nature of the actor to define a homogenous collaboration model. First, we define a new collaborative model of interaction. This model notably allows to abstract the real or virtual nature of a teammate. Moreover, we propose a new role exchange approach so that actors can swap their roles during training. The model also permits the use of physically based objects and characters animation to increase the realism of the world. Second, we design a new communicative agent model which aims at improving collaboration with other actors using dialogue to coordinate their actions and to share their knowledge. Finally, we evaluated the proposed model to estimate the resulting benefits for the users and we show this is integrated in existing CVET applications.

Keywords: Interaction for Virtual Humans, Conversational Agents, Autonomous Actors, Avatars, Virtual Reality

Introduction

The use of virtual reality for training offers lots of advantages. First, it reduces the costs and risks of the training for the trainees and the equipment. It also allows the trainees to learn

collaborative procedures, along with other team members who can either be other users or autonomous agents. Our paper focuses on such Collaborative Virtual Environments for Training (CVETs) where real users and autonomous agents efficiently collaborate toward a common goal as equal teammates. In this context, each role of the training can either be handled seamlessly by a user or an autonomous agent.

Virtual agents are crucial in trainings as they generally replace team members when there is not enough trainees. Thus, they have to be able to handle a precise role in the procedure and also help other trainees in order to enhance their learning experience. To do so, an autonomous agent should be able to (1) collaborate with other team members no matter they are users or other autonomous agents, (2) have credible behaviors and gestures to help users to comprehend its actions on its surroundings, and (3) be able to easily communicate with other team members in order to share information or simply synchronize its actions with them.

To improve the user experience as well as immersion during the training, we propose in this paper a CVET that uses such a virtual agent. In this CVET, we gathered two important contributions impacting on the collaboration between actors. Our first contribution consists in a unified interaction model for users and autonomous agents, allowing them to efficiently collaborate during the training. This model allows to abstract the nature of the actors but also to perform a role exchange between actors and to use physical interactions as well as physically-simulated

virtual agents to increase the realism of the training. Second, these agents are also communicative agents able to handle a dialogue with a user in order to furnish him details about the procedure or about the training, using its knowledge and natural language.

Related Work

Collaborative Virtual Environments for Training

In the CVET literature, autonomous agents generally interact with users in three different manners [1]: (1) as a *personal assistant* assigned to a single trainee to help him/her, (2) as a *team assistant* assuring the communication between users and helping them to coordinate their actions and (3) as an *equal team member* operating autonomously and performing the collaborative procedure alongside users and other autonomous agents. Most of the CVETs are focusing on this last case. Thus, team members have to be able to perform tasks, interact with the objects and communicate with other teammates [2]. Regardless of their nature, autonomous agents and users work towards a shared objective. In most CVETs, autonomous agents are able to perform their task independently [3, 4]. Thus, they are generally able to play different roles such as collaborators, instructors or assistants. They can also replace team members needed for a training [5, 6]. Unfortunately, interactions between team members and particularly between autonomous agents and users are limited. They perform parallel tasks, working towards the team's shared goal but cannot neither interact collaboratively on a same object nor exchange their roles during the simulation. Some recent platforms handle collaborative interactions between teammates. This is the case of the Generic Virtual Training [7, 8]. In the collaborative version of STEVE [9], agents play the double role of collaborator and instructor. As collaborators, they simply perform their part of the procedure, wait for the actions of their teammates and communicate when needed, but as instructors, they directly interact with the trainee to help in the task.

Communication with autonomous agents in CVET

Some CVET are already using conversational agents to enhance the user training user [10]. The dialogue is generally based on the knowledge of the agent and different structures are used to design this agent. For instance, an agent such as STEVE [9] uses production rules, whereas MAX [11] uses logical frames to represent the knowledge concerning the world. Both, however, use a hierarchical structure to represent the domain associated to their procedural task. Nevertheless, none of these architectures provides a unified knowledge representation allowing an actor to collaborate and communicate with other teammates to achieve a team goal.

In CVET, integrating dialogue models (state/frame based [12], agent based [13], etc.) to the task model is needed for the coordination and the sharing of the knowledge between teammates. Some dialogue systems such as TrindiKit [12] are based on Information-State (IS) updates, whereas some others such as COLLAGEN [13] are using agent based model for the conversational behavior. However, in these systems, users do not actively perform the collaborative tasks with other teammates. Furthermore, these works focused either on dialogue management for exchange of information [12] or for the planning and execution of the goal directed plan [13], but a little work [11] is done to integrate these two aspects together to achieve mutual understanding and shared goal.

Limitations

Regarding the literature, collaboration between users and autonomous agents is still limited as they barely understand their respective needs and choices. Moreover, in most CVETs, autonomous agents' actions are precomputed beforehand and not adapted on-the-fly to the users' interactions. Some solutions have been proposed to handle task-driven postures as well as physical objects interactions for autonomous characters [14, 15]. These methods, however, have not been adapted yet for CVET composed of users and autonomous agents.

In addition, the dialogues between users and autonomous agents are usually composed of

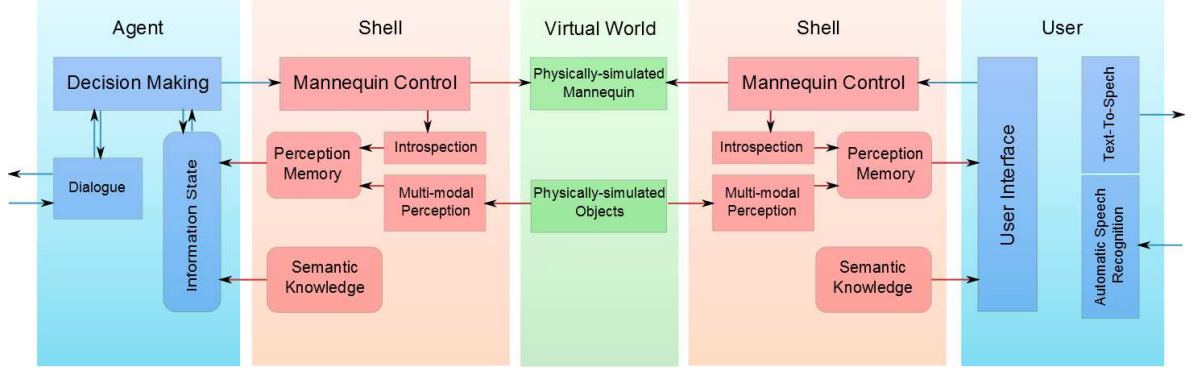


Figure 1: Global architecture of the communicative autonomous agents and users present in the virtual environment and collaborating through their Shells.

predefined sentences and are limited to instructions given by an autonomous agent. Thus, users and autonomous agents can hardly communicate with each other. Even in [16, 17], the work on conversational agents focused on a limited aspect of the dialogue and did not really considered users actively engaged in collaborative tasks with other teammates.

Overview

In order to tackle the limitations of current systems, we present a new functional CVET. Our actors' architecture is summed up in Fig. 1. To enhance the training, we integrated :

1. A model of collaborative and physical interactions. Our interaction model allows actors to act collaboratively on the same objects. Moreover, we used this model to abstract the real and virtual nature of the actors during the simulation and to allow them to exchange on-the-fly their avatars as well as the role attached to this avatar. We also extended our model to handle physically-simulated objects and avatars. Thus, movements and gestures of the autonomous agents, as well as the actors' interactions with the environment, look more natural and realistic to the users.
2. A communication module allowing trainees to communicate with autonomous agents teammates and ask them for information. Moreover, it also allows these agents to interact and collaborate with

the users using natural language. Conversational agents improve the training by updating their knowledge during the simulation and by sharing it with trainees, giving information about the procedure.

Finally, we present some results based on our concepts in the result section. First, we present two different training scenarios that answer industrial needs of our partners and that use our functionalities. Second, to demonstrate the benefits of these new functionalities, we will also present some experiments we conducted in order to evaluate our contributions. This experiments aim at evaluating three elements : (1) the exchange protocol, a concept based on our collaborative interaction model, (2) the communication between users and autonomous actors in the context of CVET using our model, and (3) the whole CVET system for novice trainees in an immersive set-up.

Collaborative Virtual Training

Contrary to real life training, the use of virtual reality offers the opportunity to learn a collaborative procedure with a single user and to use autonomous agents to replace missing members. However, current Collaborative Virtual Environments for Training (CVET) are limited since collaborative interactions on the same objects barely exist. Moreover, the collaboration between actors has to be easy and natural, even if these actors are users or autonomous agents. This way, the training can be performed even

if only a single trainee is available as the other actors can be controlled by autonomous agents. Thus, we have to define a new model for collaborative interaction as well as an abstraction of the actor similar for both users and autonomous agents.

Collaborative Interactions

Using the STORM model, presented in [18], we defined a new collaborative model of interaction. Our model, detailed in [19], allows an object to be controlled by several other objects, sharing its control with multiple sources. To do so, each object presents a set of manipulable parameters and handles itself the multiple modifications from the controlling sources. An object of the simulation can then have a new capability: either *Interactor*, if it can manipulate other objects parameters, or *Interactive*, if its parameters can be manipulated by others, or even both. In our context, an object can be a simple tool, such as a screwdriver, as well as an avatar in the environment.

Considering this, an avatar is then a specific object that can be controlled by different sources such as a user or an autonomous agent. Moreover, it could also be configured to be manipulated by multiple sources at the same time, for instance by a trainer who wants to show specific moves of a particular avatar to the trainee controlling this avatar.

The Shell

Designing a new collaborative procedure requires to acknowledge various elements concerning the users, like the knowledge they acquire and use during the procedure. While learning such a procedure, it is also important for a trainee to understand the decisions made by his teammates, which can be autonomous agents or users. In order to meet these requirements, we introduce *the Shell* as an entity containing a knowledge base and whose control can be exchanged between autonomous agents and users. We introduced the term of Shell¹ to emphasize

¹The term of *Shell* refers to the Japanese manga serie "Ghost in the Shell" written by Masamune Shirow where a "Ghost" is a thinking mind that can be embedded in a "Shell", i.e. a body, and takes control over

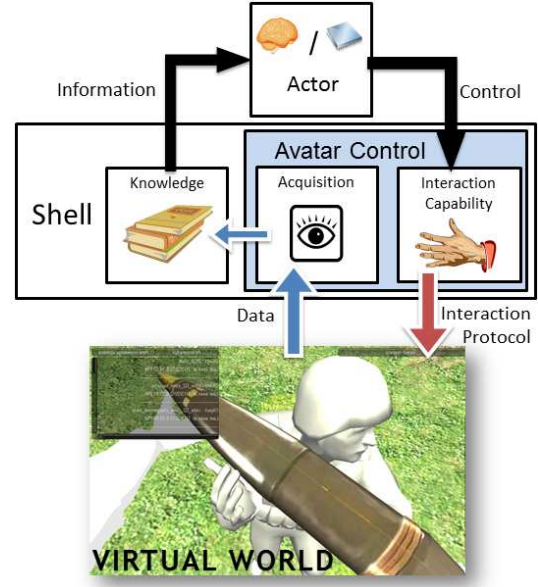


Figure 2: The Shell : This entity gathers an interaction module and a knowledge base coupled with an acquisition module.

the difference with the term of Avatar, commonly used in the Computer Graphics field, that usually designates only the visual representation (i.e. the mannequin) of an actor in the virtual world. This Shell extends the concept of avatar and allows to abstract the nature of an actor from its representation in the virtual training. Although this does not prevent an actor from having his own knowledge base, it is also essential for the Shell to save some pieces of knowledge at runtime to either allow an actor to resume its actions or to consider which knowledge has been acquired so far.

In our model, a Shell is an entity including essential components required to complete a procedure in a CVET, its architecture is presented Fig. 2. These components are the following: an interface to represent a trainee in the environment and possessing the capacity to interact with elements of the virtual world and an acquisition module to perceive information about the state of the world during the simulation, and finally a knowledge base, either gathered by the avatar, or known a priori. We find this concept to be particularly relevant in CVETs with mixed teams of

it to act in the world.

users and autonomous agents, where each actor has a role in the collaborative task, whatever its nature. Moreover, the knowledge base of a Shell can be accessed by both type of actors, for example to help in their decision process or to retrieve information about the beginning of the training. Thus, we are able to completely abstract the real or virtual nature of the actors during the task.

The Shell is designed as both an interactive object, as it is controllable by an actor, and an interactor, as it can act upon the objects of the virtual environment. Thus, using this entity, both users and autonomous agents share a common representation and common capacities in the virtual world as they both use the Shell as their entry point to interact with the virtual world.

In the context of CVET, a critical element is the knowledge gathered by the trainees at the end of the procedure. Thus, we completed the Shell's architecture by including a knowledge base. This knowledge has two main purposes. First, when a user is controlling the Shell, the knowledge is still gathered in the Shell, which helps him/her in the training as he/she can retrieve some information needed to complete the procedure. Second, when an autonomous agent is controlling the Shell, the knowledge is used to retrieve information about the previous actions of this Shell and the coming tasks, its role in the procedure, its teammates and the virtual world history to support its decision process.

To better retrieve relevant information in the Shell's knowledge, we identified different kinds of knowledge needed by a trainee to understand and carry on a procedure. First, the trainee must be able to get information about his/her avatar and its current state. Then, information concerning the procedure to perform and the individual or collaborative tasks is also needed. In the context of CVET, information about the team and the teammates is important. Indeed, actors need to be aware of their partners in order to synchronize their actions with them. Finally, the trainee must update his/her knowledge about the surroundings in the virtual world, for instance to locate needed tools or to retrieve information about them. Based on these observations, we defined four knowledge categories to sort the gathered information : Ego, Task, Team and World. The knowledge is either known a priori or filled using the acquisition module during the simula-

tion. The four categories help to easily retrieve knowledge for users, but also for decision process as well as communication for autonomous agents.

Moreover, regarding the needs of training environments, we defined an exchange protocol which provides new usages for both the trainers and the users. For instance, a teacher performing a virtual medical procedure, using a Shell, can pick a student and exchange their roles, allowing this student to perform a part of the medical procedure. In industrial training applications, a trainer could take control over an embodiment controlled by an autonomous agent and either help or perturb the work of the trainees by modifying the training conditions in real time. Using our exchange protocol, this can be done without the trainees noticing the active intervention of their teacher. This protocol mainly consists in an exchange of Shells between the two involved actors, implying an exchange of the avatars and knowledge associated to each Shell.

Physical Collaboration with Autonomous Agents

The credibility of the training is a critical issue as the more the CVET will be realistic, the better the formation will be. Indeed, the realism of the environment allows the trainee to easily convert a virtual training to a real context. Some solutions have been proposed to handle task-driven postures and physical object interactions for autonomous characters [14, 15] but have not been adapted yet for CVET. Currently, in most of the CVETs, the behavior of objects involved in the scenario, as well as the virtual teammates interactions, are precomputed and played when the trainee triggers a specific action. Thus, the training environment as well as autonomous agents do not adapt their physical behaviors accordingly with the trainees actions. Moreover, the different objects involved in the training can generally not be manipulated as freely as they would be in the real world.

In order to tackle this issue, we use in our CVET a physics simulation for both avatars and objects. In this simulation, a physically-simulated avatar is used to represent the actor. Using physics, this avatar can directly interact with the reactive virtual environment and agents

can adopt realistic postures in the virtual environment. Our implementation is based on the work of Liu et al. and concerns as well the representation of the user [20] than the representation of an autonomous agent [15].

Moreover, this avatar is controlled by the actors through the Shell which allows users and autonomous agents to have similar controls on it. In the case of an autonomous agent, the avatar controlled by the agent automatically handles the basic interactions such as "*grab an element*" or "*reach a place*" by adapting its posture in order to interact with the virtual environment. In the case of a user, many different means exist to control an avatar, from a simple clickable interface which determines the element in the focus of the user to more immersive set-ups such as tracking markers capturing the motion of the user. For non-immersive set-up, the physics simulation automatically handles the postures to adopt for the user's avatar, as it does for autonomous agents. In the case of an immersive set-up, however, the user can totally control the postures of the avatar. The use of such a physically simulated environment allows the user to perform any gesture he/she wants and to find on his/her own the preferred one to choose.

Concerning the objects of the world, the use of physics increases the realism of the interactions as their reactions correspond to real life situations. This observation is even more relevant when using an immersive set-up for the user as he/she is able to have natural interactions with the objects and have the impression to manipulate objects of different weight. He/she also has to deal with the collisions between the different elements of the environment. Thus, it facilitates the design of the environment as it directly reacts to users interactions and these interactions does not need to be scripted beforehand.

Moreover, the physics simulation could be used to check the validity of the adopted postures. To extend this contribution, it would be interesting, especially in the context of CVET, to give feedbacks to the user on these adopted postures in order to lead him/her to correct him/her if the posture is unsafe or not ergonomic [15].

Verbal Communication to Foster Collaboration

Our architecture endows the autonomous agents with both deliberative and conversational capabilities: agents are able to engage themselves into shared plans and produce the necessary dialogue acts for the coordination of their actions (Fig. 1). First, the agents' deliberative behavior is based on the joint-intention theory from [21], which states that to coordinate their activities, agents must have the joint-intention to achieve the collective goal and must agree upon a common actions plan. Second, following the principles of the shared-plan theory [22], agents make decisions and communicate in order to make commitments towards the group to achieve their common goals. The agent architecture borrows the principles of the BDI-architecture *Belief, Desire, Intention* [23, 24]. Our behavioral architecture treats deliberative and conversational behaviors uniformly as guided by the goal-directed shared activity. It considers dialogue as a collaborative activity and ensures its intertwining with the task-oriented agents activity. The last theoretical foundation of our model is the Information State approach (IS) for dialogue modeling from [12]. The IS maintains the necessary contextual information of the agent about the various dimensions of the context: dialogical, semantic, cognitive, perceptual and social.

Using these components of the architecture, the agent can combine the unified knowledge representation hosted by the Shell with its IS, in order to decide whether to elaborate the plan, react to the current situation, or exchange information with other real or virtual teammates. Therefore the agent makes decision based on the overall context of the task, its mental state and the course of the dialogue.

Knowledge Organization

Knowledge Representation. Agents deal with different sources of knowledge: information collected by the Shell along the training and sorted in four categories (Ego, World, Team, Task), semantic knowledge (static at this scale and containing information about the Task), contextual information about the ongoing decision-making and conversational processes

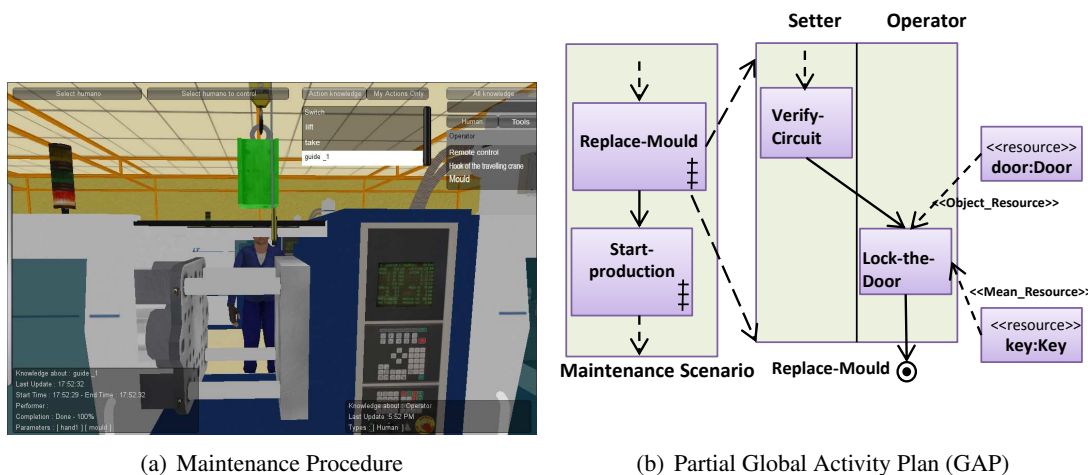


Figure 3: (a) A Setter and an Operator need to communicate and collaborate in a procedure. (b) The autonomous agent provides information about the GAP and update its knowledge.

handled by the agent.

Following the shared-plan theory, the perceived state of the current situation is an instantiation of shared concepts that agents hold in their semantic knowledge. Perception and dialogue allows agents to update their knowledge using the Shell, and to monitor the progress of the shared activity. This condition ensures the consistency of the agent's knowledge after any update of its belief initiated by communication or perception. Besides, although the agents share the same semantic knowledge, due to their location and perception, they only have partial beliefs about the world and do not necessarily share the same information about the situation.

Information about the task is the central element of the agents' knowledge base. It defines the global activity plan (GAP) and the conversation protocols (see below), which are represented as hierarchical frame structures. On the one side, the GAP is shared by all actors and each of its node represents sub-goals that must be achieved by the team. A sub-activity plan, modeling a shared activity, is represented as a node of the GAP that refers another graph. A quick overview of this GAP is shown in Fig. 3. On the other side, the conversation protocols are used to make the dialogue evolve towards the achievement of common goal and they can be viewed as sets of production rules.

Information-State. The agent architecture integrates *deliberation* and task-oriented *communication* aspects together by using the semantic data structure *Information-State*. It contains contextual information of dialogue such as conversation context or social context. The social context also includes an *agenda* of dialogue goals. These contextual features are inherited from the model defined in [25] that we extended by introducing the new component *Task Context*. It contains agent’s desires, goals related to the task and task intentions. The agent uses this *Task Context* to monitor the task in progress and to provide this context to the conversation. Thus, agents are able to talk about the ongoing task, the properties of the objects in the environment, their own intentions and the shared goals.

Collaborative and Conversational Behavior

Decision-Making. Our conversational agents have to make decisions about their course of actions, their knowledge update and the necessity to exchange information with others. The decision-making mechanism fulfills these two functions: deliberative control and belief revision. The deliberative control aims at deciding which goal the agent must pursue. It uses semantic information about the GAP, the IS and the knowledge from the Shell to decide



(a) Military training example



(b) Industrial training example

Figure 4: Collaborative Virtual environments developed using the presented concepts.

whether a conversational behavior is required to collaborate with other team members. Our belief revision maintains the consistency of both the knowledge base and the IS by updating the agent’s beliefs using (1) its current perceptions through the Shell about resources and capabilities and (2) the new information acquired after a Shell exchange with another actor.

The decision-making algorithm verifies whether the agenda in IS is not empty or holds some communicative intentions. If so, control is passed to the conversational behavior. Otherwise, the agent chooses the plan to be executed using a deliberative BDI-style behavior. If the agent identifies some cooperative situations in the collective activity, then the control passes again to the conversational behavior: the cooperative situations generate communicative intentions in the agenda, which causes the agent to exchange information with other team members. Then, the control is passed to the deliberative behavior which generates new intentions. Finally, the agent selects actions to execute and updates its IS to maintain the knowledge about the current task context.

Conversational Behavior. Our agents hold reactive and proactive conversational abilities. Reactive behavior allows them to manage information-seeking dialogues. Thus, users can ask questions to their virtual partners about the ongoing activity (*e.g.* actions of the procedure, resources, current goal, state of objects). Proactive behavior corresponds to so-called delibera-

tion dialogues. Here, the agent generates its own communicative intentions for the coordination of the collaborative activity. We defined three collaborative communication protocols that lead the agent to engage itself into this kind of dialogue ??:

- (1) When the agent decides to pursue a new collective goal, it communicates with other team members to establish joint-commitment, and to ensure that everyone is going to use the same plan of action to achieve this goal.
- (2) When the agent has performed its planned actions and the shared activity is not yet finished, it requests information to other team members to know when the activity will be finished.
- (3) The agent, who has just finished the last action of the shared activity, informs other team members that the activity has ended.

Results

Two Industrial Applications

Using the concepts presented in sections and , we developed two different CVETs. These two environments have common needs. First, they both involve several trainees that are controlled seamlessly by users or autonomous agents with no impact on the learning procedure. Second, they can be used on various virtual reality platforms, from the computer station to a complete immersive system using tracking devices. And finally, the use of physic and communication in both of these trainings is an asset improving the learning of the user.

Military application The first virtual environment for training we designed is a collaborative procedure involving five trainees. The procedure consists in the deployment of a military vehicle. An overview of the resulting environment is shown in Fig.4(a).

Industrial application The second scenario is a virtual factory where two teammates have to learn how to exchange the mold of an industrial machine. This environment is visible Fig.4(b).

Moreover, a video demonstrating each of our contributions (independently as well as together in the same application) is available at <http://youtu.be/5olmgpxoTUg>.

Evaluation of the Theoretical Concepts

In order to measure the benefits of our CVET, we designed three experiments. The first experiment aimed at evaluating the metaphors used in the exchange of Shell in order to determine how such a metaphor should be defined for a better comprehension and use. This experiment and the results are detailed in [26]. The conclusion of this experiment is that different parameters should be taken into account depending on the application context (user-friendly versus efficiency) but also that different metaphors should be used regarding the familiarity of the end-user with virtual reality applications which is really important in the design of CVET.

Moreover, we conducted two other experiments. The first experiment concerned the verbal communication between the user and the autonomous virtual agents in the context of a collaborative task. This experiment mainly aimed at analyzing how the user's activity articulates with those of the virtual teammates. The preliminary results, based on verbalizing analysis, highlights that users consistently react to the information provided by the agents and to their questions or requests.

Finally our last experiment aims at evaluating the collaborative training system with end-users. Its purpose is to evaluate the 'usability' (ISO 9241-11) of the global system. This evaluation consists in a collaborative scenario with five actors where one of them is a real user and have to collaborate, synchronize and dialog with four

autonomous agents in order to complete a complex collaborative task and to correct their mistakes. This experiment is currently on-going.

Conclusion

We presented a whole and complex model of CVET intensively using autonomous agents to assist trainees as full-fledged team members. Our contributions come under two main areas. First, we presented an advanced model of collaborative interactions. This model has been completed by the Shell, to gather the control of an avatar and its associated knowledge and to abstract the real or virtual nature of the actors. Moreover, we proposed to increase the user experience during the training using a physics simulation for the avatars and the objects of the virtual world. This contribution aims at intensifying the realism of the feedbacks given to the user and allows him to interact more naturally with the environment. Thus, it can be seen as a baseline of reasoning components that can be considered when bidding new CVETs.

Second, we detailed a new model of communicative agent that handles specific constraints of CVET. Thanks to these communicative behavior, agents are now able to engage themselves into shared plans and to produce and interpret dialogue acts to accomplish commitments about their achievement towards collaborative goals and for the coordination of their actions with users and other autonomous agents. Moreover, they are able to share they knowledge on the procedure or on the elements of the simulated environment with other trainees, which is a real benefits for users as they can ask them for help as they would do with a real teammate.

To conclude, we presented a model of autonomous agents that perfectly fits the various needs of CVET. This agent is able to collaborate seamlessly with virtual agents as well as users. This collaboration is possible at different level from the collaborative actions to a collaborative conversation and knowledge exchange. Not to mention the precious asset of the use of physics which helps to reinforce the user's immersion during the collaboration. This whole CVET has been developed in the french research project CORVETTE involving academical and

industrial partners. More details on the project and on the different individual contributions can be found on the CORVETTE's website². Our current research now aims at evaluating the concept of our entire CVET with end users in order to estimate the benefits of our various contributions. We will also estimate these benefits regarding the set-up used for the training, from the basic computer station to the entire immersive set-up.

Acknowledgements

This work was supported by the French Research National Agency project named CORVETTE (ANR-10-CONTINT-CORD-012) and by the French Unique Interdepartmental Funds SIFORAS (FUI 11).

References

- [1] K. Sycara and G. Sukthankar. Literature review of teamwork models. Technical report, Robotics Institute, 2006.
- [2] W. R. Swartout, J. Gratch, R. W. Hill, E. H. Hovy, S. Marsella, J. Rickel, and D. R. Traum. Toward virtual humans. *AI Magazine*, 27(2):96–108, 2006.
- [3] J. Dugdale, B. Pavard, N. Pallamin, M. el Jed, and C. L. Maugan. Emergency fire incident training in a virtual world. In *Proceedings ISCRAM2004*, 2004.
- [4] L. Edward, D. Lourdeaux, J.-P. Barthes, D. Lenne, and J.-M. Burkhardt. Modelling autonomous virtual agent behaviours in a virtual environment for risk. *International Journal of Virtual Reality*, 2008.
- [5] P. Chevaillier, T.-H. Trinh, M. Barange, P. De Loor, F. Devillers, J. Soler, and R. Querrec. Semantic modeling of virtual environments using mascaret. In *Software Engineering and Architectures for Real-time Interactive Systems*, 2012.
- [6] C. Barot, D. Lourdeaux, J.-M. Burkhardt, K. Amokrane, and D. Lenne. V3S: A virtual environment for risk-management training based on human-activity models. *Presence: Teleoperators and Virtual Environments*, 22(1):1–19, 2013.
- [7] N. Mollet and B. Arnaldi. Storytelling in virtual reality for training. In *Edutainment*, pages 334–347, Hangzhou Chine, 2006.
- [8] S. Gerbaud, N. Mollet, F. Ganier, B. Arnaldi, and J. Tisseau. GVT: a platform to create virtual environments for procedural training. In *IEEE VR*, 2008.
- [9] J. Rickel and W. L. Johnson. Virtual humans for team training in virtual reality. *Applied Artificial Intelligence*, 1999.
- [10] B. Lok. Teaching communication skills with virtual humans. *IEEE Computer Graphics and Applications*, 2006.
- [11] N. Leßmann, S. Kopp, and I. Wachsmuth. *Situated Interaction with a Virtual Human - Perception, Action, Cognition*, pages 287–323. Walter de Gruyter, Berlin, 2006.
- [12] D. R. Traum and S. Larsson. The information state approach to dialogue management. In *Current and New Directions in Discourse and Dialogue*, Text, Speech and Language Technology. 2003.
- [13] C. Rich, C. L. Sidner, and N. Lesh. Collagen: applying collaborative discourse theory to human-computer interaction. *AI Magazine*, 22(4):15–25, 2001.
- [14] K. Yamane, J. J. Kuffner, and J. K. Hodgins. Synthesizing animations of human manipulation tasks. In *TOG*. ACM, 2004.
- [15] M. Liu, A. Micaelli, P. Evrard, and A. Escande. Task-driven posture optimization for virtual characters. In *ACM SIGGRAPH/Eurographics Symposium on Computer Animation (SCA)*, 2012.
- [16] D. Traum, S. Marsella, J. Gratch, J. Lee, and A. Hartholt. Multi-party, multi-issue, multi-strategy negotiation for multimodal virtual agents. In *Intelligent Virtual Agents*, 2008.

²<http://corvette.irisa.fr>

- [17] H. Buschmeier and S. Kopp. Towards conversational agents that attend to and adapt to communicative user feedback. In *Intelligent Virtual Agents*, 2011.
- [18] N. Mollet, S. Gerbaud, and B. Arnaldi. STORM: a generic interaction and behavioral model for 3D objects and humanoids in a virtual environment. In *IPT-EGVE*, 2007.
- [19] Andrés Saraos Luna, Valérie Gouranton, and Bruno Arnaldi. Collaborative Virtual Environments For Training: A Unified Interaction Model For Real Humans And Virtual Humans. In *Edutainment*, pages 1–12, 2012.
- [20] M. Liu, A. Micaelli, P. Evrard, A. Escande, and C. Andriot. Interactive dynamics and balance of a virtual character during manipulation tasks. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, pages 1676 – 1682, Shanghai, China, May 2011.
- [21] P. R. Cohen and H. J. Levesque. Confirmations and joint action. In *Proceedings of IJCAI’91*, pages 951–957, 1991.
- [22] B. J. Grosz and S. Kraus. Collaborative plans for complex group action. *Artificial Intelligence*, 86(2):269–357, 1996.
- [23] A. S. Rao and M. P. Georgeff. BDI agents: From theory to practice. In *First international conference on multi-agent systems*, pages 312–319, 1995.
- [24] V. Mascardi, D. Demergasso, and D. Ancona. Languages for programming BDI-style agents: an overview. In *Proceedings of WOA’05*, 2005.
- [25] Harry Bunt. Multifunctionality in dialogue. *Computer Speech & Language*, 25(2):222–245, 2011.
- [26] T. Lopez, R. Bouville, E. Loup-escande, F. Nouviale, V. Gouranton, and B. Arnaldi. Exchange of avatars: Toward a better perception and understanding. 2014.